4. Consistency Checks

4.1.   Interdependencies Between Components of Retirement Income

4.1.1. Objectives

Task 3 is concerned with both internal and external model consistency. Internal consistency refers to corrections for correlation among income components that are projected separately. External consistency refers to consistency of summary statistics[27] from model projections with external macroeconomic or other models that project similar summary statistics. This task applies to both our own projections and those of The Urban Institute and Brookings Institution.

Subtask 3-1 is to identify the components of each individual's retirement income that are modeled separately, assess whether they should be consistent with each other, and recommend techniques to make them consistent.

4.1.2. Overview

A number of potential behavioral and/or correlational interactions exist between the various components of retirement income. By far the most important issue for any policy simulation exercise is that potential behavioral responses to change may not be reflected in projections of retirement income taken one at a time and assuming exogeneity of other sources of income, for example. The modeling and estimation of the relevant behavioral relationships are the subject of current debate and should be the subject of further research. They are outside the scope of the current effort but should be kept in mind and improved whenever possible. We focus here on a discussion of more correlational interactions among sources of retirement income, but even these are not fully understood and may be outside the scope of the current effort. Our purpose is to raise the issues for discussion and potential resolution or postponement for later analysis.

4.1.3. Discussion

A number of sources of positive correlation exist between various sources of retirement income, which, if ignored and if each source of retirement income is treated as independent, will potentially bias analyses of the distribution of retirement income, incidence of poverty, and like calculations.

---

[27] There is one more important consistency, namely between the SIPP (and/or PSID) and the U.S. population. While both surveys aim to be representative of the population, discrepancies between survey aggregates and U.S. aggregates may arise from the way questions are asked and other sources. It is the responsibility of The Urban Institute/Brookings Institution and ourselves to adjust parameter estimates for such discrepancies before applying them in the projections. We anchor the demographic transition models to vital statistics, as described in detail in Panis and Lillard (1996).

Retirement income is strongly affected by, if not largely determined by, pre-retirement behaviors and outcomes, which may induce correlation in sources of retirement incomes. And post-retirement behavior, such as saving or the lack thereof, may reflect similar pre-retirement behavioral patterns before.

Dimensions of pre-retirement behavior and outcomes that affect retirement income include: (1) life cycle patterns of work and earnings; (2) marriage, divorce, and the work and earnings of spouses; (3) work-related fringe benefits, including health insurance and DC and/or DB pension accumulation; (4) Social Security contributions and benefits; and (5) pre-retirement saving and financial asset/wealth accumulation.

Because these pre-retirement outcomes are obviously important determinants of financial resources — assets and income — in retirement, important interactions among pre-retirement outcomes will affect the distribution of retirement income. While the following discussion raises a number of potentially important interactions from the literature, the current state of knowledge and agreement on these issues does not permit easy incorporation into the MINT modeling. Instead, it suggests important issues for further research, and potentially feasible adjustments to the current effort may be forthcoming from discussion of the issues as work progresses.

There are a number of relevant relationships from the literature (published and not yet published). An obvious first relationship is that both Social Security and pension benefits are related to the level and pattern of pre-retirement life cycle earnings. Lillard and Weiss (1997), for example, report a positive relationship between post-retirement income from Social Security and from pensions using the old Longitudinal Retirement History Survey data. Hurd, Lillard and Panis (1998) show that workers with smaller retirement accounts are more likely to cash-out the account and spend the money when changing employers near retirement. In addition, some DB pensions are reduced if the beneficiary receives Social Security benefits. The SIPP does not offer this type of information; in the Health and Retirement Study (HRS), about 15 percent of respondents indicated that their future private pension would be offset by Social Security benefit receipts. The opposite may also apply: A Social Security benefit payable to a (divorced/surviving) spouse may be reduced if the person receives a periodic payment based on his or her own employment that was not covered under Social Security from the Federal Government, a State, or a political subdivision of a State (§ 1836, Social Security Handbook 1997).

Marital status is clearly important, and not only because of the spouse and survivor benefits of Social Security. Smith (1997) reports greater assets and asset accumulation in the PSID for persons married over a five-year period than for unmarried persons over the same period. Unpublished results by Lillard and Karoly (1997) show that for men, household wealth over the life cycle is strongly positively related to their own permanent earnings and only weakly related to the permanent earnings of the women to whom they are married. On the other hand, for women, household wealth over the life cycle is strongly positively related to the permanent earnings of the men to whom they are married and only weakly related to their own

earnings. As a result, retirement assets and income from assets will be positively related to marital history.

The level of saving and the form of saving are clearly important. Pre-retirement and post-retirement consumption and saving behavior are likely to closely related, so that those persons who do not save for retirement are least likely to maintain assets after retirement, and there are substantial proportions of the population with little or no assets at retirement. Similarly, those persons with little or no savings at retirement are likely to have low wage rates, so continuing to work is a relatively unattractive option and their Social Security replacement rates are high, as noted by Hubbard, Skinner and Zeldes (1997). The form of saving may also be important. For example, 401(k) assets may be offset by mortgage debt (Engen and Gale 1997). However, there is some controversy around that issue. See, for example, Poterba, Venti, and Wise (1995); Venti and Wise (1996); and Poterba, Venti, and Wise (1996).

There also may be interactions between work and Social Security or pension benefits. Social Security benefits are reduced if the beneficiary's earnings exceed a certain threshold. For example, beneficiaries age 65-69 may earn $14,500 in 1998; for each $3 in additional earnings, their benefits are reduced by $1. While Social Security benefits may be reduced based on the beneficiary's earnings, they may be taxed based on any income. Social Security benefits are partially subject to income taxation for higher-income retirees. Beneficiaries with incomes of more than the base amount ($25,000 if single and $32,000 if married) are liable for income taxes on a portion of their OASI benefits. This correlation applies to after-tax income (Task 4), but also has behavioral implications, as in the next item. Partial taxation of Social Security benefits on the basis of other income components (e.g., from assets) may reduce incentives for labor force participation and thus earnings among the elderly (Part I, Task 4). Conversely, a reduction in Social Security benefits (which may be simulated as part of a proposed policy change) may induce higher labor force participation and earnings among the elderly.

4.2. Identify Macroeconomic Models and Their Use for Benchmarking

4.2.1. Introduction

This subtask identifies appropriate macroeconomic forecasting models for validating, and possibly benchmarking, the MINT microsimulation model.[28] The goal of this subtask is to support analyses of the extent to which aggregate forecasts implicit in MINT are consistent with accepted, external forecasts.

Linking micro- and macromodels at some level is fairly common.[29] For example, a non-central part of a microsimulation model may rely on simple time-series techniques to forecast a necessary macro-level variable. Alternatively, a macroeconomic model may be fully imbedded into a microsimulation model so that behavioral feedbacks between the models may exist. For this subtask, we focus on a third type of link that uses a macroeconomic forecasting model to validate the implicit aggregate forecasts of the microsimulation model.[30] If the forecasts are close by some metric, then the microsimulation model is considered "valid."

We proceed as follows. Section 4.2.2 describes the characteristics of the ideal macroeconomic forecasting model for validating MINT, without concern for whether such an ideal actually exists. We conduct a survey of macroeconomic forecasting models and select the most promising candidate models in Section 4.2.3. Section 4.2.4 discusses the mechanics of aggregating, validating, and benchmarking MINT.

4.2.2. Macroeconomic Model Characteristics

Before we examine potential macroeconomic models, we first list the characteristics of a macroeconomic model that would be appropriate for validating MINT. It should not be expected that a macroeconomic forecasting model will contain all these characteristics; rather, these criteria will serve as a guide in the selection of an appropriate macroeconomic forecasting model.

The ideal macroeconomic model for our purposes possesses the following characteristics. The model must:

- *Forecast through 2020.* MINT forecasts retirement income through the year 2020. A macroeconomic model that serves to validate MINT projections must therefore be specified to forecast through at least the year 2020.

---

[28] We gratefully acknowledge substantial expert input from Steven Haider.
[29] Anderson (1990) provides a detailed discussion of different types of linkages with numerous examples.
[30] Sargent (1985) provides a detailed discussion of general validating procedures for simulation models.

- *Have been evaluated for accuracy.* The purpose of the validation procedure is to compare MINT to macroeconomic forecasts that are considered "good." The macroeconomic model must therefore have a proven track record, for example, through comparisons of past projections with actual outcomes.
- *Forecast different income types.* Most macroeconomic models forecast aggregate quantities such as GDP. By contrast, MINT will project individual income components as well as aggregate retirement income. Ideally, the macroeconomic model should forecast the same types of personal income, such as income from earnings, pensions, Social Security, savings, etc.
- *Forecast by cohort.* MINT is specified to forecast retirement income of the 1931-1960 birth cohort only. The ideal macroeconomic model therefore forecasts income by birth cohort, permitting a direct comparison.
- *Provide confidence intervals.* A formal statistical test of whether two forecasts differ requires confidence intervals (i.e., standard errors) for both forecasts.
- *Be fully documented.* Projections of the macroeconomic model and MINT may only be expected to be identical if their underlying assumptions are the same. Where identical assumptions are lacking (as will very often be the case), discrepancies may arise. An evaluation of the source of such discrepancies requires full knowledge of both models' underlying assumptions.
- *Be in the public domain or owned by SSA.* Many macroeconomic forecasting models are maintained by private, for-profit consulting firms. These firms tend to charge considerable sums for the use of their models, which in itself is an argument for preferring public-domain models. More importantly, a model that is in the public domain or owned by SSA is more likely to permit a thorough examination of the underlying structure than proprietary models. It may furthermore be more readily adapted to serve MINT's needs.
- *Support projections under alternative policy regimes.* MINT's purpose is to permit scenario analysis of alternative policy reforms. The ideal macroeconomic model is capable of forecasting under these same alternative policies, so that both the baseline and the alternative projections may be validated. Preferably, no assistance from the model's keeper is required to project under alternative regimes.
- *Include behavioral feedbacks.* To ensure the feasibility of developing a microsimulation model in a timely fashion, the MINT architects have accepted a minimum number of behavioral feedbacks. The ideal macroeconomic model includes behavioral feedbacks, so that it supports future extensions of MINT and generates a rough indication of the bias introduced by the lack of behavioral feedbacks in MINT.

Although we do not think it is useful to explicitly rank the importance of each of these characteristics, we do consider it useful to indicate which characteristics are most critical. Specifically, an appropriate macroeconomic model for validating MINT should at least (a) forecast through 2020, (b) have been evaluated for accuracy, and (c) be fully documented.

4.2.3. *Survey of Forecasting Models*

Types of Macroeconomic Forecasting Models

Building on Brayton et al. (1997), we distinguish three types of forecasting models.[31] The first type of forecasting model is the traditional "IS-LM" (or Keynesian) macroeconomic model. These large-scale models specify the relationship between the macro-level variables, often explicitly including sluggish price adjustment. IS-LM models have been quite successful at forecasting the quarter-to-quarter performance of the economy. Examples of these types of models are the Fair Model (maintained by Ray C. Fair at Yale University) and the Washington University Macro Model (maintained by Macroeconomic Analysts).

The second type of forecasting model are "modern macro" models or "small scale macro models."[32] These models are characterized by the inclusion of optimizing agents with expectations that are explicit and rational. They tend to exclude certain sectors of the economy to remain tractable, and often feature only a limited number of policy levers. Although these models do not tend to forecast short-term fluctuations in the economy as well as IS-LM models, they are well equipped to forecast the long-term impact of policy regime changes.

The final type of forecasting model are largely statistical models. Such models fully capture the dynamic relationship among a few variables by relying on vector auto-regressive (VAR) techniques for estimation and projection. However, these models include few, if any, economic relationships. They tend to forecast the short-term quite well but are unable to forecast the effects of policy changes.

These three categories are not mutually exclusive. Many forecasting models will rely on model characteristics from different categories, depending on the particular forecasting goals. For example, the Federal Reserve Board's forecasting model for the United States (FRB/US) includes an underlying IS-LM model for the macro economy and an explicit characterization of expectations. Furthermore, many forecasting models will use a largely statistical model for aspects of the economy that are secondary to the modeling goals.

For purposes of MINT validation, the most promising models belong to the IS-LM class. They are most often used for forecasting, including by the Federal Reserve Board, the Congressional Budget Office, and the Council of Economic Advisers. They tend to perform quite well. However, the models tend to have at least one major drawback (in addition to often being proprietary): They usually include a "judgmental adjustment" or "fudge factor" that is to a large extent arbitrary and precludes the calculation of standard errors.

---

[31] See Fair (1994) for a historical review of macroeconomic forecasting models.

[32] See Leeper and Sims (1994) for a particular small-scale modern macro model. Also, real business cycle (RBC) models are considered small-scale modern macro models.

Potential Macroeconomic Forecasting Models

We surveyed the following macroeconomic models.

- BC: The Blue Chip Financial Forecasts are a consensus forecast of leading private economists and analysts. They forecast major financial characteristics of the United States such as GDP, Federal Funds Rate, and the Prime Rate. More information may be found on their web-site (http://www.bluechippubs.com).
- CBO: The Congressional Budget Office forecasts are used for federal governmental activities such as revenue and deficit planning. The CBO 10-year forecasts are readily available on the web (http://www.cbo.gov/reports.html). The CBO generates its forecasts both from its own analysis as well as using the forecasts from other models (DRI, MA, BlueChip and WEFA). See Congressional Budget Office (1998) for more details.
- COREMOD: This model uses the WUMM (see below) as its basis for short-run forecasts and uses a neoclassical growth model with representative households and firms (and myopic expectations) for long-run forecasts. The model was developed and is maintained by Macroeconomic Associates, LLC (http://macroadvisers.com); the forecasts are provided on a fee-for-service basis.
- DRI: This forecasting service is run by Standard and Poor (http://www.dri.mcgraw-hill.com). Five-year forecasts are publicly available.
- FAIR: The Fair model was developed by Ray C. Fair at Yale University. The model is a large-scale IS-LM model. Notably, the model is extensively documented, in the public domain, and available on-line for the public to generate its own forecasts (http://fairmodel.econ.yale.edu). See Fair (1994) for a detailed description of the model.
- FRB/US: The Federal Reserve Board model for the U.S. was developed at the Federal Reserve Board in the mid-1990s. It is a large-scale macro model that relies on an IS-LM framework for short run behavior, and it includes explicit expectations and optimizing agents for longer-run behavior. See Brayton et al. (1997) for a historical perspective on the development of the model and Brayton and Tinsley (eds., 1996) for a more technical description of the FRB/US model.
- MDM: The Macroeconomic-Demographic Model was developed as a microsimulation model usable for policy analysis by the National Institute on Aging in the 1980s. The MDM completely integrates a macroeconomic forecasting model developed by Dale Jorgenson and Edward Hudson. The model is capable of generating forecasts at least 75 years into the future and was used for extensive policy analysis during the 1980s. See National Institute on Aging (1984) for further details.
- RSQE: The RSQE model is maintained at the University of Michigan. It generates forecasts for 3 years. More information is available on its website (http://rsqe.econ.lsa.umich.edu).
- SSA: The Social Security Administration generates economic and demographic forecasts that it regularly uses to evaluate the long-run solvency of its various programs. For an example of its forecasts, as well as for information on how they are generated, see Board of Trustees (1997).

- <u>WEFA</u>: This group was formerly the Wharton Econometric Forecasting Associated established at the University of Pennsylvania by Lawrence Klein. The group merged with Chase Econometrics in 1987 to become WEFA; see its website (http://www.wefa.com). It provides forecasts for up to 25 years. The forecasts are generated with a traditional IS-LM, large-scale model.
- <u>WUMM</u>: The Washington University Macro Model was developed at Washington University and is now maintained by Macroeconomic Associates, LLC (http://macroadvisers.com). Its forecasts are provided on a fee-for-service basis. The model is a traditional IS-LM model that generates forecasts for 10 years.

Table 4.1 summarizes the basic characteristics of the major forecasting models that we examined.

Table 4.1. Forecasting Models

Model Name and Source	Acronym	Fully documented?[a]	Public domain?	Projection horizon
Blue Chip Consensus Forecasts	BC		No	5 years
Congressional Budget Office	CBO		Yes	10 years
COREMOD, Macroeconomic Advisors	CORE-MOD	Yes	No	>50 years
Standard and Poor' s DRI	DRI		No[b]	5 years
Fair Model, Ray C. Fair at Yale University	FAIR	Yes	Yes	5 years
Federal Reserve Board, U.S. Quarterly Model	FRB/US	Yes	Yes[c]	>50 years
Macroeconomic-Demographic Model	MDM	Yes	Yes	>50 years
RSQE, University of Michigan	RSQE		No[b]	3 years
Social Security Administration Forecasts	SSA	Yes	Yes	80 years
Wharton Econometric Forecasting Associated	WEFA		No[b]	25 years
Washington University Macro Model, Macroeconomic Advisors	WUMM	Yes	No	10 years?

Notes: [a] We only mark "yes" for models for which we have acquired and examined detailed documentation. Such documentation may be available for other models, but because of undesirable model attributes (such as too short a forecast period), we have not acquired the documentation.
[b] Although some aspects of the forecasts are publicly available, the forecasts are compiled by a private company on a fee-for-service basis. Complete information will likely have to be purchased.
[c] Technically, this model is in the public domain. Realistically, assistance will be required from staff members at the Federal Reserve Board.

The table does not include a column with an indication of accuracy of historical model predictions. All models claim to be accurate, though few back it up in the documentation to which we have access. In addition, most evaluation period are very short (typically six quarters or less), even for models that forecast over long periods.

Recommendation

Obviously, none of the models satisfies all desired criteria as outlined in Section 4.2.2.

The COREMOD model appears reasonably well-suited for validation purposes. Among its disadvantages are that it (as many other IS-LM based models) includes a fudge factor to arrive at "desirable" predictions, and that it is not in the public domain.

The FAIR model does not include a fudge factor, is superbly documented, fully in the public domain, and executable on the Internet. Its main disadvantage is that it only projects out five years into the future. It is not always clear to us what criteria underlie the number of years that models chose to project out. With today's computing resources, there should not be any technical reasons. IS-LM models that provide long-run projections tend to use the IS-LM structure for a limited number of years and adopt broad trends thereafter. Perhaps the FAIR model may be adapted to support longer projection periods, but this would require assistance from Ray C. Fair, its architect.

The FRB/US model does not use a fudge factor to arrive at desirable predictions and it does provide standard errors of its predictions. Our conversations with Federal Reserve Board staff members indicated that the FRB is eager to see external application of its models, and willing to provide assistance.

The SSA projections are not based on an explicit model. SSA considers projections generated by other models and bases its own projections on the OASDI Trust Funds Board's best estimate of the future course of the population and the economy. It provides three projections termed "intermediate," "low cost," and "high cost." The intermediate projections represent the Trustees' consensus expectation of moderate economic growth through the projection period.

Based on the minimum requirements of forecast horizon through 2020, proven accuracy, and extensive documentation, we recommend the FRB/US and COREMOD models. In addition, we recommend comparisons with the SSA forecasts to ensure consistency across projections developed by SSA actuaries and MINT.

4.2.4. Validating and Benchmarking MINT

In this section, we address the mechanics of validating and, if necessary, benchmarking MINT. This section applies to the use of macroeconomic models and their use for benchmarking, so we will restrict the discussion to income flow projections. Section 4.3 treats validation of demographic projections using macro models that are not economic in nature.

We first describe issues in aggregating MINT's micro forecasts to the national level. Second, we draw attention to the MINT sample universe and contrast it to universes to which macroeconomic models apply. Third, we discuss the statistical and heuristic validation of MINT results vis-à-vis macroeconomic forecasts. Finally, we suggest adjustments that may be made to MINT projections to correct for discrepancies from macroeconomic forecasts.

Aggregation

In principle, aggregation of micro projections to the national level is achieved by simply computing a weighted sum:

$$Y^M = \sum_i w_i y_i,$$

where y_i is a projected vector of outcomes for individual i (such as of various income components or total retirement income), w_i is the (scalar) weight for individual i, and Y^M is the aggregated outcome vector. Aggregation of microsimulation projections to the national level thus requires proper weights, w_i.

MINT is based on the 1990, 1991, 1992, and 1993 SIPP panels. Projections are created only for individuals born in the years 1931-1960, and only for full panel respondents, i.e., only for those who responded to all waves of a particular panel. The Census Bureau provides weights for these simulants such that the weights for each panel add up to the covered sample universe.

The fact that MINT is based on multiple SIPP waves slightly complicates the weighting procedure, since untransformed weights would result in a weighted population of about four times the actual sample universe. Since each panel is designed to be representative of its sample universe, the normalization procedure makes no difference in expectation of Y^M, only in the efficiency of the projection (variance of the estimate Y^M). Simply dividing all weights by four to account for four panels yields an unbiased but inefficient estimate of aggregate income flows. The most efficient projection is obtained by weighting proportional to sample size in the projection. Specifically, let n^j denote the number of respondents in SIPP panel j (j=90,…,93). The most efficient transformed weights are:

$$w_{ji}^* = \frac{n^j}{n^{90} + n^{91} + n^{92} + n^{93}} w_{ji},$$

where w_{ji} is the full panel weight for individual i in SIPP panel j. Table 4.2 shows sample sizes for the four SIPP panels and the corresponding most efficient weights. The 1991 SIPP panel was smaller than the others and its weights (pnlwgt) are largest, on average, so that the efficient weight factor is smallest. The table only includes individuals with positive full panel weight, pnlwgt, and only individuals born in 1931-1960. For sample selection details, see Section 2.6.

Table 4.2. Simulation Sample Sizes and Efficient Weight Factors

SIPP panel	sample size	weight factor
1990	16,821	0.2825
1991	11,914	0.2001
1992	15,491	0.2602
1993	15,311	0.2572
Total	59,537	1.0000

Comparability

MINT projections are based on the SIPP (with certain exclusions), which is representative of the U.S. civilian noninstitutionalized population as of the SIPP survey years. In accordance with the definition of Gross Domestic Product (total output produced within the borders of a country), macroeconomic models generally apply to all residents of United States territory. There are thus several discrepancies between the SIPP sample universe and the sample upon which most macroeconomic models are based. These discrepancies are likely to lead to discrepancies in projections.

- In 1990, approximately 600,000 individuals were in military quarters (1990 Census). These are excluded from the SIPP, which covers the civilian population only, but included in the universe underlying macroeconomic models.
- In 1990, approximately 3.3 million individuals were institutionalized. Of these, 1.75 million were in nursing homes, presumably not many from among the 1931-1960 birth cohorts. However, 1.13 million were incarcerated and another 0.44 million were in mental hospitals, juvenile institutions, or other institutions (1990 Census). The institutionalized population is excluded from the SIPP, but included in many macroeconomic models.
- MINT projections only account for individuals living in the United States as of the SIPP baseline survey. They thus exclude income from immigrants that enter the United States between the baseline survey years and 2020. During the 1980s, the annual gross inflow of legal immigrants was around 1.0 million, and the annual net inflow around 0.8 million (McCarthy and Vernez, 1997). Future flows are, of course, highly dependent on immigration policy. The Immigration Act of 1990 substantially increased the number of legal immigrants permitted starting in 1992.

Each of these discrepancies points at population counts and thus income flows that are smaller in MINT than in macroeconomic models. Over a 25-30-year period, the largest discrepancy probably stems from immigration. Its magnitude is difficult to ascertain, since assumptions on immigration policy in most macroeconomic models are not explicitly specified. An exception is SSA projections, which assume total annual net legal immigration to rise to 900,000 by the year 2000 and remain constant thereafter (intermediate scenario). COREMOD simply features an exogenous parameter for total annual population growth (without distinguishing birth and immigration).

Two additional comparability issues arise. First, ideally, we would like to separately validate MINT projections of income components (from partial labor force participation, from pensions, from assets, etc.), but to our knowledge, no macroeconomic model provides projections of income components. For practical purposes, validation should thus be restricted to MINT's aggregate retirement income projections.

Second, since MINT only takes the 1931-1960 birth cohorts into account, aggregate MINT projections cannot be directly compared to macroeconomic model projections, which do not support projections by cohort. A first order solution inflates aggregate MINT projections for the 1931-60 cohorts to be reflective of the entire population. To avoid problems stemming from the changing age distribution, this is best done by computing the income position of the elderly relative to the general population as of the SIPP survey years and by multiplying the projections by the inverse of that relative income. A more sophisticated inflation procedure would account for temporal changes in the relative income position of the elderly.

Validation

Many statistical tests are available for testing the equality of two vectors of outcomes, such as one generated by MINT and the other by a macroeconomic model. Consider the following test statistic:

$$Q = \left(Y^M - Y^E\right)' V^{-1} \left(Y^M - Y^E\right),$$

where Y^M is a vector of outcomes generated by MINT, Y^E a vector generated by the external macroeconomic model, and V the covariance matrix of their difference. The vector of outcomes may represent, for example, various income components at a point in time, total retirement income at multiple future points in time, or a combination thereof. If $\left(Y^M - Y^E\right)$ is normally distributed, Q asymptotically follows the c^2 distribution with a number of degrees equal to the dimension of the outcome vector.[33] Suppose that only one (scalar) outcome is compared to an external model. This special case reduces asymptotically to a simple z-test, with the variance given by:

$$V = s_{Y^M}^2 + s_{Y^E}^2 - 2s_{Y^M Y^E}.$$

In practice, this formal statistical approach is unlikely to be feasible. The variance of Y^M is very difficult to determine, especially if MINT's projections need to be scaled up to reflect all cohorts; the covariance between Y^M and Y^E is unknown and probably not zero.

In practice, a more heuristic approach must therefore be taken. A "small" discrepancy, defined by the MINT user, may be acceptable, especially if differences in underlying assumptions and sample universes provide an explanation for the direction and magnitude of the discrepancy. A "large" discrepancy, however, should be analyzed carefully and may lead to the discovery of modeling errors.

Benchmarking

As a matter of principle, we do not advocate adjusting MINT model parameters to ensure a close match between aggregate MINT projections and external

---

[33] It may be preferable to formulate the test statistic in terms of the logarithm of outcomes.

macroeconomic projections. Discrepancies may arise from many sources, including from imperfections in macroeconomic models. However, we acknowledge the desirability of comparability of simulations conducted using MINT and other models, which is facilitated if the baseline simulations are equal.

If changes to MINT are contemplated to ensure a match of the aggregate projections of MINT and of an external macroeconomic model, we strongly recommend that every effort be made to preserve substantive conclusions from running MINT. For example, the calibration should not affect measures of income inequality such as the Gini coefficient. The following procedures may preserve MINT's conclusions.

- Adjust all weights proportionally. This may force equality of future income flows at the expense of a realistic current income flow.
- Adjust an intercept and/or a slope equally for all respondents. This may preserve current income flows and generate future flows that match those of an external macroeconomic model. While these types of adjustments ("fudge factors") are common in many macroeconomic models, we offer this option with great reluctance. The adjustment is arbitrary and the model is no longer capable of generating confidence intervals.

As a general rule, adjustments should be made equally for all simulants to prevent building in distributional effects that MINT itself does not generate.

4.3. Identify Demographic Models for Validating MINT

4.3.1. Introduction

This section identifies appropriate demographic forecasting models for validating, and possibly benchmarking, the MINT microsimulation model. The eventual objective is to support analyses of the extent to which aggregate forecasts implicit in MINT are consistent with external forecasts.

The process of validating and benchmarking a microsimulation model with demographic forecasts is exceedingly similar to the process with macroeconomic forecasts; thus, much of the discussion in Section 4.2 is applicable to validation with demographic forecasts. Rather than repeat criteria listed above, we consider this section as an extension of the previous one.

We proceed as follows. Section 4.3.2 outlines the ideal characteristics of a useful demographic forecasting model for validating MINT. We provide a brief review of demographic forecasting models and a discussion of the most promising models in Section 4.3.3. Further information about the theory and mechanics of aggregating, validating, and benchmarking MINT was discussed in Section 4.2 and will not be repeated here.

4.3.2. Demographic Model Characteristics

Before we turn to candidate demographic models, we first list the characteristics of a demographic model that would be appropriate for validating MINT. The ideal demographic model for evaluating MINT should:

- *Forecast through the year 2020.* MINT implicitly forecasts the population through the year 2020. A demographic model that serves to validate MINT projections should thus forecast through at least the year 2020.
- *Have been evaluated for accuracy.* The purpose of the validation procedure is to compare MINT to demographic forecasts that are considered "good." The external demographic model should therefore have a proven track record for accuracy. However, it is much less common to validate demographic models than it is to validate macroeconomic models. See Keyfitz (1981) for an exception.
- *Forecast population size and demographic characteristics by cohort.* MINT is specified to forecast retirement income of the 1931-1960 birth cohort only. Suitable demographic models should therefore forecast changes in the population by cohort, permitting a direct comparison.
- *Provide confidence intervals.* A formal statistical test of whether two forecasts differ requires confidence intervals (i.e., standard errors) for both forecasts.

Moreover, for policy analysis, standard errors are necessary to evaluate the potential risk under various scenarios.[34]

- *Be fully documented.* Projections of the demographic model and MINT may only be expected to be identical if their underlying assumptions are the same. Where identical assumptions are lacking (as will very often be the case), discrepancies may arise. An evaluation of the source of such discrepancies requires full knowledge of both models' underlying assumptions.

We limit our consideration to forecasting models that contain all these characteristics, with the exception of confidence intervals.[35]

4.3.3. *Survey of Forecasting Models*

Although there is long history of forecasting population trends, there are many fewer models in existence that are used for forecasting. We first provide a brief review of types of forecasting models that are generally used for demographic forecasting; then, we specifically discuss candidate models. [36, 37]

There are three primary categories of demographic forecasting methods. The first category is demographic accounting methods. These methods specify accounting identities for the underlying components of population growth. Then, these underlying components are forecasted into the future. For example, the population in a given year can be defined as the population in the previous year plus births last year and minus deaths last year (ignoring migration for this simple example). Births and deaths are forecast individually and population growth follows from the accounting identity. Usually, the underlying components are forecast using so-called "informed judgment methods." For example, the U.S. Census Bureau uses informed judgment methods. (See U.S. Bureau of the Census (1984) for a detailed description of its forecasts.) Although these models tend to provide reasonable long-term forecasts, standard errors are not available. To address the issue of variance, forecasts are sometimes made under "low," "intermediate," and "high" assumed levels for key

---

[34] Tuljapurkar (1992) provides a succinct introduction about why it is difficult to obtain standard errors with demographic forecasts and provides a discussion of the importance of standard errors for demographic forecasts in a decision theoretic framework.

[35] Additionally, one could require that the models are in the public domain, support projections under alternative policy regimes, and include behavioral feedbacks. These criteria, however, play a less important role for demographic models than they do for macroeconomic models. All demographic models that we encountered are in the public domain. Moreover, while changes to Social Security policy may theoretically have some effect on mortality, marriage, and divorce rates, these effects are likely to be small and perhaps ambiguous or controversial. In our assessment, these criteria do not merit further attention for purposes of evaluation and benchmarking MINT.

[36] Candidate models that are not reviewed here include DYNAMISM and MDM. Both models are older and do not focus on demographic forecasting. For more information on both models, see Gordon and Michel (1980) and the National Institute on Aging (1984).

[37] For a useful review of forecasting methodologies, see Land (1986). We borrow heavily from Land for this section. For a review that focuses on forecasting for older populations, see Guralnik, Yanagishita, and Schneider (1988).

trends, but the choice of such levels is typically arbitrary and the value of the additional projections difficult to assess. [38]

The second category is time series methods. These methods rely on projecting demographic quantities using standard time series models such as an AutoRegressive Integrated Moving Average (ARIMA) model. The models are purely statistical without underlying demographic justification. They tend to predict short- and medium-term population changes quite well.

The third category methods are structural in nature. Many macroeconomic models include endogenous population characteristics such as fertility, mortality, and marriage. The structural models are well equipped to forecast potential demographic changes to policy changes, but there is not a clear consensus that the added complexity is necessary to achieve sensible results for population forecasting.

Potential Demographic Forecasting Models

While there is a rich literature on models of mortality, marriage formation, and marriage dissolution, few models have been used to project population counts of the United States, by cohort and marital status. The best known and most widely cited population forecasts are those produced by the Bureau of the Census in its P25 Series of publications. Report P25-1130 projects population by age, sex, and race/ethnicity. The Census Bureau itself does not project population counts by marital status. In specialized reports, such as Bureau of the Census (1996, P23-190), population projections by marital status are provided. Those projections, however, are not produced by the Bureau of the Census, but by Felicitie Bell of the Office of the Chief Actuary of the Social Security Administration. The latest such report is Actuarial Study No. 112, Bell (1997).

Bell (1997) uses a demographic method based on accounting identities. Her projections include components for fertility, mortality, immigration, marriage, and divorce. [39] Following the practice of the U.S. Bureau of the Census, Bell uses informed judgment methods for each component series. Her assumptions, however, do not always correspond to those of the Bureau of the Census.

---

[38] Alho and Spencer (1990) provide a useful discussion of the difficulties of interpreting the high and low scenarios generated as part of the forecast.

[39] Bell (1997) projects for the Social Security Area, consisting of residents of the 50 states and the District of Columbia; armed forces overseas; civilian residents of Puerto Rico, the Virgin Islands, Guam, American Samoa, Palau, and the Northern Mariana Islands; Federal civilian employees overseas; dependents of Armed Forces and Federal employees overseas; crew members of merchant vessels; and other citizens overseas. The SIPP is representative of the civilian non-institutionalized population within the 50 states and the District of Columbia. Any comparison of MINT with Bell (1997) should thus adjust for military personnel, institutionalized persons, and persons outside the 50 states and the District of Columbia.

An important difference is in assumed reductions of mortality risks. Bell reports that the average annual percentage reductions in age-adjusted central death rates between 1900 and 1994 are 0.94 for men and 1.33 for women. In recent years, gains in longevity have been smaller: 0.78 percent annually for men between 1982 and 1994, and 0.54 percent for women. Bell's intermediate assumption is that male and female mortality rates will both decrease at an average rate of 0.56 percent per year during the period 1994 through 2071. These reductions are far smaller than the average annual reduction observed between 1900 and 1994, about one-sixth below the reduction over the past twelve years, and in fact lower than any period of this century, with the exception of the 1954-1968 period. As a result of this assumption, life expectancy at birth is assumed to increase from 72.6 in 1995 to 77.5 in 2050 for males, and for females from 79.0 in 1995 to 82.9 in 2050.

The Census Bureau makes complicated adjustments to male and female age-specific death rates, resulting in life expectancy for males and females in 2050 of 79.7 and 84.3 years, respectively (P23-1130, Table B-1). The implicit reductions in mortality rates of the Census Bureau are thus substantially larger than those of Bell (1997).

Our MINT mortality projections assume annual reductions in mortality rates of 0.81 percent for males and 1.41 percent for females. These reductions are based on hazard model estimates of pooled Vital Statistics age-specific death rates from 1901 to 1994. Only mortality rates of individuals age 30 or over are taken into account in those estimates, as this is the relevant age group for MINT purposes.

Bell's assumptions on marriage and divorce are based on trends in marriage and divorce rates over the past 15 years or so. Marriage rates are based on data from the Marriage Registration Area (MRA), which comprises 42 states and the District of Columbia, and covers approximately 80 percent of all marriages in the United States. As shown in Figure 4.1, marriage rates have declined mildly since about 1970. Bell (1997) assumed for the intermediate alternative that future age-adjusted rates of marriage for the Social Security Area would continue to slowly decrease and then stabilize in 2021. She inflated the number of MRA marriages to reflect the entire United States, but then reduced it by 5 percent to correct for Nevada; Nevada is not in the MRA, but has disproportionally many marriages.

Bell's divorce rates are based on the Divorce Registration Area (DRA), which comprises 31 states and covers approximately 48 percent of all divorces. As shown in Figure 4.2, the divorce rate increased substantially in the 1960s and 1970s, but leveled off after 1980. Bell (1997) assumed under the intermediate alternative that throughout the projection period the age-adjusted divorce rate would remain close to the level as recently experienced.

MINT's assumptions are based on hazard model estimates of marriage and divorce experiences reported by SIPP respondents in the Marital History Topical Module and subsequent panel reports until the end of the survey waves. For male marriage rates, we found a reduction of 0.79 percent annually; for females, we found a reduction of

0.36 percent annually (Table 2.5 on page 17). Male divorce rates were found to be flat since 1980; based on marriage histories reported by women, divorce rates have creeped up by 0.58 percent annually since 1980 (Table 2.7 on page 29).[40] We assume that those historical trends will continue through the projection period.

---

[40] Data limitations required estimating male and female divorce models separately, even though conceptually they follow the same process. See Section 2.4.
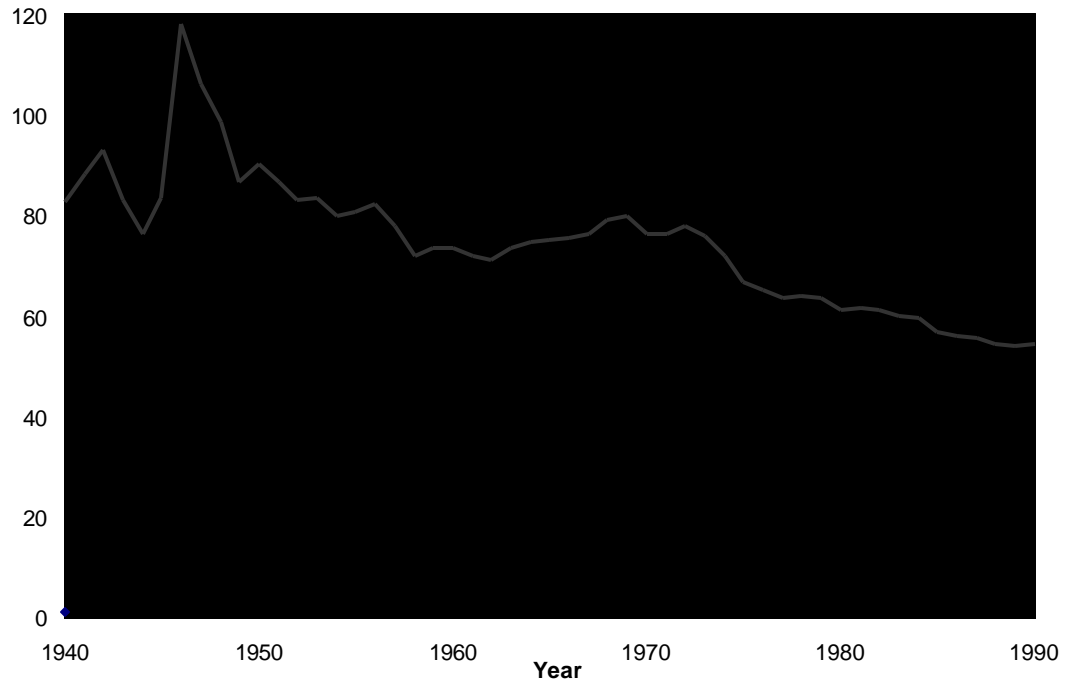
Figure 4.1. Marriages per 1,000 Unmarried Women Aged 15+, 1940-1990
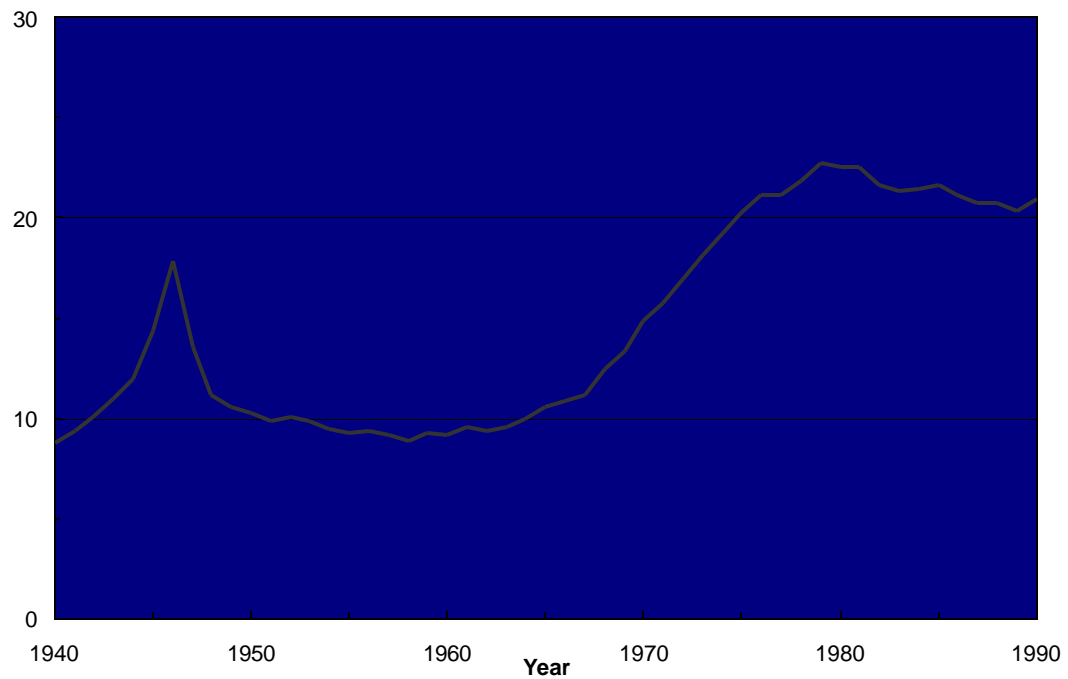

Figure 4.2. Divorce Rate per 1,000 Married Women Aged 15+, 1940-1990

Section 4.5 compares MINT's projections by marital status to Bell (1997).

Both Bell (1997) and the Bureau of the Census (1996b) use "informed judgment methods" to obtain assumptions on future trends in demographic submodels. Both also formulate "low-cost," "intermediate," and "high-cost" scenarios. A drawback to this approach is that standard errors are not available for the forecasts and that the assumed high/low ranges for births, deaths, and other demographic components are not probabilistically consistent with one another. Furthermore, the choice of low-cost and high-cost scenarios is largely arbitrary and their corresponding projections difficult to interpret.

Ronald Lee and Shripad Tuljapurkar address this issue in a series of papers. See, for example, Lee and Tuljapurkar (1994). The Lee-Tuljapurkar approach uses a combination of demographic models of fertility and mortality and a statistical time series model. In the first step, simple demographic models of age- and time-specific fertility and mortality are estimated, resulting in estimates of time trends with standard errors (much like our estimates of the mortality, marriage formation, and divorce time trends from Vital Statistics and SIPP). The estimates are translated into stochastic transition matrices. In the second step, most recent population counts (by sex and age) are multiplied by the transition matrices to obtain next-period population counts (with a correction for assumed immigration). The number of projection period years determines the number of matrix multiplications. This stochastic method enables Lee and Tuljapurkar to generate standard errors and confidence bands for their forecasts.

The Lee and Tuljapurkar (1994) expected total population projection for the year 2065 corresponds closely to the Census Bureau's intermediate projection, and their 95 percent confidence intervals are close to the low and high scenario projections. However, Census low/high intervals for elderly subpopulations are much wider than those found using stochastic forecasts, whereas the opposite was found for youth and elderly dependency ratios. [41]

We believe that the Lee-Tuljapurkar method offers a far superior alternative to the low/high scenarios used by Bell (1997), the Bureau of the Census (1996b), and many other "informed judgment" approaches. It is internally consistent, statistically sound, and not subject to arbitrary assumptions (except on exogenous influences, such as immigration).

Recommendation

An attractive feature of Bell (1997) is that it projects population counts by marital status. However, we question Bell's intermediate assumption of 0.56 percent annual

---

[41] Alho and Spencer (1990) attempted to relate high/low forecasts of the Office of the Actuary to confidence intervals. They found that the intervals may not be interpreted as confidence intervals, and that the high/low intervals tend to be wider than the 95 percent confidence bands.

mortality rate reductions. Comparisons of MINT population forecasts with Bell (1997) result in larger MINT projections because of MINT's assumption of more rapid gains in longevity; see Section 4.4. Additional discrepancies may result from differences in the population under consideration. In particular, Bell's inclusion of the military and the incarcerated or otherwise institutionalized population is likely to exacerbate the discrepancy. [42]

Projections of the Bureau of the Census (1996b, Current Population Series P23-1130) and Lee-Tuljapurkar (1994) yield aggregate results that are very close to each other. The Census low/high intervals, however, differ from Lee-Tuljapurkar's 95 percent confidence bands. Given the arbitrary nature of the Census Bureau's high/low assumptions and the internal consistency of the Lee-Tuljapurkar method, we favor benchmarking MINT's aggregate population projections with those of Lee and Tuljapurkar (1994).

Lee and Tuljapurkar (1994) project expected aggregate population size in 2020 of 316 million; the number of individuals age 65+ is projected at 52 million. MINT is restricted to the 1931-1960 birth cohort, excludes immigrants, the military, incarcerated and other institutionalized individuals, and should thus project smaller population counts.

Neither Lee and Tuljapurkar (1994) nor the Bureau of the Census (1996b) provide estimates of population counts by marital status. We therefore recommend that MINT projections by marital status be benchmarked against the results in Bell (1997). As argued above, we question Bell's mortality assumptions, but her projected shares of individuals that are never married, married, widowed, and divorced may serve as useful benchmarks. Table 4.3 shows the projected fractions by age group for the year 2020.[43]

Table 4.3. Projected Marital Status of Persons Age 65 and Over by Sex, 2020 (Bureau of the Census 1996a)

	Males				Females			
	Single	Married	Widowed	Divorced	Single	Married	Widowed	Divorced
Age 65+	6.2	72.1	12.7	8.9	5.0	43.6	37.1	14.3
Age 65-74	7.5	75.0	7.6	9.9	5.6	55.4	22.4	16.6
Age 75+	4.0	66.9	22.1	7.0	4.3	28.3	56.2	11.2

---

[42] Bell's high-cost alternative assumes mortality rate reductions about the same as for 1900 through 1994. That assumption, however, is combined with high-cost marriage and divorce rates.

[43] These figures are based on Bureau of the Census (1996a), Table 6-1, and provide greater detail than Bell (1997).

4.4. MINT vs SSA OACT Longevity Projections

Our mortality projections are based on hazard model estimates from PSID data that are corrected such that, in the aggregate, mortality rates are identical to those based on Vital Statistics data. See Section 2.2.3. Vital Statistics from 1901 through 1994 indicate that the log-hazard of mortality has decreased at an annual rate of 0.81 percent (males) and 1.41 percent (females). A key assumption underlying our mortality projections is that mortality rates will continue to improve at this pace.

Table 4.4. Historical Average Annual Percentage Reductions in Age-Adjusted Central Death Rates (Bell, 1997)

	1900-36	1936-54	1954-68	1968-82	1982-94	1900-94
Male						
0-14	2.91	4.75	1.66	4.39	2.60	3.26
15-64	1.02	1.91	-.20	2.22	.61	1.14
65-84	.20	1.15	-.13	1.47	1.21	.65
85+	.22	1.21	-.89	1.56	-.34	.38
65+	.20	1.16	-.33	1.49	.79	.58
Total	.78	1.60	-.21	1.78	.78	.94
Female						
0-14	3.12	5.01	1.72	4.19	2.49	3.36
15-64	1.19	3.62	.57	2.20	.70	1.66
65-84	.36	2.06	1.07	2.01	.58	1.07
85+	.23	1.21	.13	2.06	.09	.66
65+	.32	1.82	.77	2.03	.42	.95
Total	.90	2.47	.77	2.15	.54	1.33

Based on Table 4.4, SSA OACT observes the following (Bell 1997):

> An examination of the age-adjusted death rates since 1900 reveals several distinct periods of mortality reduction. During the period 1900 to 1936, annual mortality reduction averaged about 0.8 percent for males and 0.9 percent for females. Following this was a period of rapid reduction, 1936 to 1954, in which mortality decreased an average of 1.6 percent per year for males and 2.5 percent for females. The period 1954 to 1968 saw an actual increase for males of 0.2 percent per year and a much slower reduction of 0.8 percent per year for females. From 1968 through 1982 rapid reduction in mortality resumed, averaging 1.8 percent for males and 2.2 percent for females, annually. From 1982 to 1994, slower reduction in mortality resumed, decreasing an average of 0.8 percent for males and 0.5 percent for females.

After reviewing cause-specific mortality rates, Bell (1997) makes the following assumption:

> After adjustment for changes in the age and sex distribution of the population, the intermediate alternative mortality is projected to decrease at an average rate of 0.56 percent per year during the period 1994 through 2071, about half the average annual reduction observed during 1900 through 1994, but greater than the female rate of reduction for the 1982 through 1994 period.

As may be expected, our assumed 0.81 percent (males) and 1.41 percent (females) annual decrease in mortality rates implies greater projected gains in longevity than those based on the 0.56 percent (males and females) assumed by SSA OACT.

The first row in Table 4.5 shows remaining life expectancies for a 65-year-old person, by sex and year, as generated by the projection algorithms of OACT and MINT.

Table 4.5. Remaining Life Expectancies at Age 65 by Sex and Year

	Male		Female	
Year	OACT	MINT	OACT	MINT
1995	15.6	15.2	19.0	19.5
2005	16.0	15.8	19.5	20.6
2015	16.4	16.5	19.8	21.8
2025	16.8	17.1	20.2	22.9
2035	17.3	17.7	20.7	24.1

The 1995 OACT figures are actual figures from Vital Statistics. Note that MINT implies a slightly lower life expectancy for men in 1995 and a slightly higher one for females. This is because MINT's mortality rates follow from a fitted model; the 1995 actual life expectancy for men was above the trend, while for women, it was below the trend.

As expected, MINT projects faster gains in longevity than OACT. Between 1995 and 2035, MINT projects gain for men of 2.5 years, while OACT projects gains of only 1.7 years. For women, the difference is larger: MINT 4.6 years vs. OACT 1.7 years.

It should be noted that MINT distinguishes a male and a female time trend, and it assumes that the 1901-1994 trends will continue. These trends imply a continued and further widening of female longevity advantage. Over the recent past, we find no evidence of male-female convergence, and we therefore project a continued divergence for the next twenty-five years or so. In the long run, however, male and female trends may converge or not diverge further. We are therefore hesitant to project much beyond the MINT horizon of 2020.

4.4.1. "Current" and "Cohort" Life Expectancies

Publications such as the *Vital Statistics of the United States* series (e.g., National Center for Health Statistics, 1998) contain so-called current, "snapshot," or "cross-sectional" lifetables, which report age-specific mortality rates of the population over the period of interest. The many mortality rates in such lifetables are often summarized in a life expectancy figure. This life expectancy is based on a synthetic cohort of individuals and its computation assumes that this cohort is subject throughout its existence to the age-specific mortality rates observed for an actual population during a particular period. For example, the 1995 life expectancy at birth (75.8 years) assumes that someone who is born in 1995 will face the same mortality risks at age ten as a ten-year-old in 1995, and the same mortality risks at age 60 as a 60-year-old in 1995, etc. The MINT life expectancies reported in Table 4.5 are computed using projected "current" mortality rates as of the years indicated in the first column. While not explicitly stated, the OACT figures are also based on "current" mortality rates; the 1995 life expectancies are virtually identical to those reported in *Vital Statistics of the United States, 1995* (National Center for Health Statistics, 1998).

However, longevity has steadily increased over the past century as a result of improved nutrition, health habits, medical technology, etc. If such improvements continue in the future, a 65-year-old in 2060 is likely to experience far more favorable survival chances than a 65-year-old in 1995. The average lifespan of all children born in 1995 is thus likely to be greater than the current life expectancy of 75.8 years. How much greater depends on the rate at which mortality risks decrease.

Given the important role that trends in mortality risks play in MINT projections, we constructed "cohort" lifetables and life expectancies, also known as "longitudinal" and "generational" lifetables (National Center for Health Statistics, 1997). These cohort life expectancies are based on projected mortality rates, taking into account lower future mortality rates, as projected by the model. In other words, the remaining life expectancy for a 65-year-old in 1995 is computed using the mortality rate that was actually experienced by 65-year-olds in 1995,[44] the mortality rate that is projected for a 66-year-old in 1996, the rate projected for a 67-year-old in 1997, the rate projected for a 68-year-old in 1998, etc.

Figure 4.3 shows a stylized model of cohort mortality rates that is not based on any empirical estimates; it only serves to illustrate the difference between current and cohort rates. The top line represents the hypothetical mortality rate (in logarithmic form) of the population in 1995; slightly below it is the hypothetical pattern of the 1996 population; through the 2015 population. As shown in the figure, mortality rates are assumed to decrease steadily over time. Consistent with our mortality model, all mortality rates decrease with the same annual percentage reduction, i.e.,

---

[44] The MINT mortality model is based on 1901-1994 Vital Statistics, so strictly speaking, the 1995 rate for 65-year-olds is a projection.

the mortality patterns are parallel in logarithmic form. Following the 65-year-old in 1995 to older age and future years, we note that the effective mortality risks that he will experience (denoted by the darker line) increase less steeply with age than those of any current mortality risk pattern.
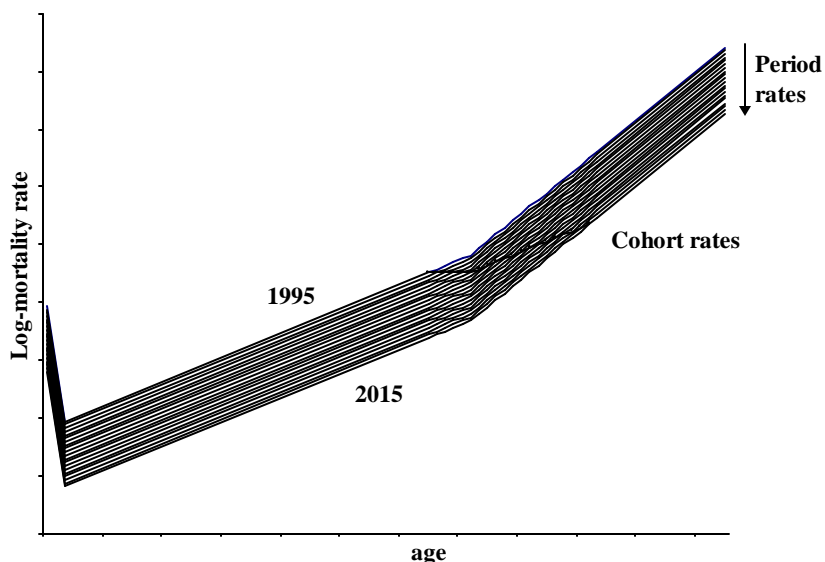

Figure 4.3. Stylized Model of Cohort Mortality Rates

The difference in current and cohort life expectancies is substantial, as Table 4.6 shows. In 1995, 65-year-old men may expect to live about 0.8 years longer than standard, current lifetables indicate. The difference is even larger for women, who may expect to benefit about 1.8 years from future developments that prolong life. The difference is larger for women because of their faster mortality rate reductions: 1.41 percent annually compared to 0.81 percent annually for men.

Table 4.6. Current and Cohort Remaining Life Expectancies at Age 65 (MINT)

	Male		Female	
Year	Current	Cohort	Current	Cohort
1995	15.2	16.0	19.5	21.3
2005	15.8	16.6	20.6	22.6
2015	16.5	17.3	21.8	23.9
2025	17.1	18.0	22.9	25.2
2035	17.7	18.7	24.1	26.6

For comparisons of life expectancies projected by MINT and external demographic models, any consistently defined life expectancy measure will suffice. Given the practice of SSA' s OACT (and the Bureau of the Census, 1996a, 1996b), we presented such comparisons using current life expectancies (Table 4.5 above). However, such life expectancies substantially underestimate the number of years that an individual

may expect to live, with potentially serious implications for the timing of retirement, savings behavior, etc. For such purposes, cohort lifetables and cohort life expectancies are superior. The Urban Institute converted wealth stocks into retirement income flows using an annuitization algorithm that RAND developed (Toder et al., 1999, pages 62 and 201); that algorithm uses cohort mortality rates.

4.5. MINT vs. SSA OACT Projections by Marital Status and Sex

We now turn to a comparison of MINT and OACT population distribution forecasts by marital status and sex. The OACT forecast method is documented in Bell (1997), *Social Security Area Population Projections: 1997*. That publication, however, does not contain full details by age group. We therefore also rely on Table 6-1 of Bureau of the Census (1996a), *65+ in the United States* (Current Population Reports, P23-190). That Census Bureau report is based on OACT's *Social Security Area Population Projections*.

As discussed earlier, OACT assumes a 0.56 percent annual mortality reduction and thus projects a smaller population size than MINT. For purposes of the current comparison, we ignore the number of deceased persons and focus on the distribution by sex and marital status only. The comparison refers to January 1, 2020. OACT projections cover all birth cohorts; MINT projections only cover individuals born in 1931-1960 birth cohort (59-88-year-olds in 2020).

OACT generates three projections: an intermediate, low-cost, and high-cost scenario. The low-cost and high-cost assumptions are chosen by consensus opinion about the plausible ranges of the forecasted series. They lack statistical basis and may not be interpreted as confidence intervals. For comparison purposes, we therefore restrict ourselves to the intermediate forecasts.

Table 4.7. Demographic Distribution in 2020, ages 65+ (percent)

	Males		Females	
	MINT	OACT	MINT	OACT
Never married	4.7	6.2	5.6	5.0
Married	76.3	72.1	46.3	43.6
Widowed	6.5	12.7	29.5	37.1
Divorced	12.5	8.9	18.6	14.3
Total	100.0	100.0	100.0	100.0

Table 4.7 shows the distribution of marital status by sex as generated by MINT and by OACT's intermediate forecasts. Several discrepancies deserve attention. First, MINT projects a lower fraction of never married men. This may in part be attributed to MINT's mortality model, which accounts for differential survival of never married males. As Table 2.1 on page 17 indicates, never married males experience mortality rates that are about 21 percent higher than those experienced by married men. The resulting shorter life expectancy implies that disproportionally many never married men will have become deceased by 2020.

Second, OACT projects higher widowhood rates in 2020 than MINT. This may in part be attributable to OACT's conservative assumption about future gains in longevity. As a result, OACT projects higher mortality and thus higher widowhood rates. Another factor is the full 65+ age range covered by OACT projections; MINT

projections only apply to individuals up to age 88 in 2020 (the 1931 birth cohort).
The 89+ population contains disproportionally many widows, as shown in Figure
2.10.

Third, MINT projects somewhat higher fractions of married individuals, a necessary
implication of lower mortality and widowhood rates. The higher projected number of
married couples implies higher Social Security expenses on spousal benefits.

Fourth, MINT projects higher fractions of divorced individuals. As shown in Section
2.8, about 61 percent of divorced women at age 62 were married more than ten years
and thus potentially eligible for benefits on the basis of their ex-husband' s earnings.
MINT may thus project greater outlays on spousal benefits than OACT, depending on
the assumptions OACT makes about the fraction divorcées that receives spousal
benefits.

Table 4.8. Demographic Distribution in 2020, ages 75+ (percent)

	Males		Females	
	MINT	OACT	MINT	OACT
Never married	3.5	4.0	4.3	4.3
Married	78.4	66.9	37.2	28.3
Widowed	7.4	22.1	42.0	56.2
Divorced	10.8	7.0	16.5	11.2
Total	100.0	100.0	100.0	100.0

Table 4.8 shows the distribution of individuals age 75 and older in 2020, by marital
status and sex. Elderly men are projected to be predominantly married. However,
their numbers have thinned, resulting in an increased fraction of widowed women.
The discrepancies between MINT and OACT projections of individuals age 65 and
older (Table 4.7) persist and become more pronounced at ages 75 and older in Table
4.8.